

module_prob4_lecture6

Fri, 2/18 12:51PM 7:22

SUMMARY KEYWORDS

cdf, slope, probability, probability density function, pdf, minus h, likelihood, definition, recall, concept, calculus class, point, continuous random variable, onwards, smaller, similar, flat, adding, starts, occurring

SPEAKERS

Sumon Majumdar



Sumon Majumdar 00:24

So now to get an idea of how can we look at which points are more likely and which are less likely, I'm going to introduce a new concept called the probability density function, or the PDF. Now, if you think back and recall from when we had discrete random variables, and we had CDFs, we could use them to get the probability of a particular point by looking at the difference in the CDF between two subsequent points, right. So the idea was this, that if I had a point X one, and I had another point X zero. So if I looked at the CDF at x zero, that was the summation of all the probabilities up to this point, if I looked at f of x one, that's the summation of all the probabilities up to this point. So if we subtract f of x one minus f of x zero, I could get exactly the probability at this point. So we're going to use a similar concept also for continuous random variables. Now, here, I've taken a continuous random variable X , and I've drawn its CDF. So it starts off at zero, and it ends up at the point. So what we want to know is, what are the in all of these points in between, right, let's say a point like x one, what's the relative likelihood of it occur? So we want to use a similar concept that we did in the discrete case. But here, of course, there is no $x \dots x$ zero right next to x one. So we're going to borrow a concept from calculus. So what do we need to look at is a nearby point, say x one minus h ? And what we're going to look at is how much does the CDF change between these two points? So what's the rate of change of the CDF between the points X one minus h and x one? So if you, if you recall your calculus class, right, so the rate of change is given by f of x one minus f of x one minus h over h . So this is the change in the y over the change in x . And now, if I take this point x one minus h , ever so close to x one, and how do I do that? I take the limit as h goes to zero. So if I do this, then this gives me the slope of the CDF at this point. And what the slope tells you is how much are we adding in terms of likelihood between the points x one minus h and x one? How much are we adding to the CDF, right between these two points? If the slope is flat, that means there's a lot of probability being added between these two points. In other words, the probability that or the likelihood of this point occurring is pretty high. On the other hand, if the slope is low, right, so for example, in this case, right here, the slope is very flat. That means between two subsequent points, not much probabilities getting out at all. So the likelihood of of these values occurring is pretty low. Right? So if we look at the slope of the CDF at a particular point, that gives us an idea of the relative likelihood of whether this point occurring is high, or whether it's low. And this particular definition is exactly the definition of the probability density

function. So the probability density function we usually denoted by $f(x)$ that's small f of x . And this is the slope of the CDF, $F(x)$ at the point x . It's the derivative of F of the CDF at that particular point is the density function. And technically, this is given by the definition, by this definition, the difference between $F(x)$ and $F(x - h)$ divided by h as h goes to zero. But if you recall your calculus class, this exactly is the definition of the derivative of the of a function. So it's the derivative of the CDF at that particular point, that gives the PDF small f of x . So in this graph here, what I've done is here is my CDF. And here is my PDF. So what I've done is that look, if I look at the CDF, and we look at the slope here at zero, right, so that means in this range, the PDF is zero, here, the slope starts increasing, here, it becomes really high, right? So the PDF here, it goes up above that at this point onwards, right, the slope starts becoming flatter and flatter, and ultimately it becomes zero. So the density here, it becomes smaller and smaller and smaller. And ultimately, here it becomes. And what this PDF tells me is that these points are relatively more likely to occur, this point has almost the highest likelihood of occurrence. But here onwards, again, the probabilities start falling, and here the provable the likelihood of those points occurring becomes. So this is the definition of the probability density function. So let me stop the clip here. This concept is a relatively hard concept. So I would ask you to take a minute to think about this. And in the next clip, I'll come back and apply to an example and show you how you can actually derive a PDF for us